

生成AIと著作権： NYTとOpenAIの訴訟を参考に

2023年12月、NYTが自社の著作権を侵害されたとしてOpenAIとマイクロソフトを訴えた。生成AIによる著作権侵害は今後大きな問題となるだろう。日本でも生成AIによる著作権侵害へどのように対処すべきかの議論が活発化している。金融機関も自社の著作物をいかに保護すべきかを具体的に検討すべきだ。

NYTがOpenAIを提訴

2023年12月27日、米ニューヨーク・タイムズ社（以下、NYT）は、生成AI大手OpenAIと、OpenAIの主要な出資元であるマイクロソフトを、自社のコンテンツに対する著作権侵害で提訴した¹⁾。NYTは、OpenAIのAIモデルが無断で数百万件の自社コンテンツを学習データとして利用し、その結果自社の利益およびブランドを毀損していると主張している。請求された賠償額などは公開されていないが、NYTは訴状の中で「記事データの無断利用による損害は数十億ドルに上る」としている。

原告のNYTの主張は多岐にわたるが、大別すると、一つは「コンテンツへのタダ乗りによる経済的損失」と、もう一つが「OpenAIが生成するコンテンツによるNYTの信頼性の毀損」の二つに集約できるようだ。

前者に関してNYTはOpenAIが提供するChatGPTが過去のNYTの記事のコピーとも見える回答を出力していることを示し、さらに後者ではNYTのコンテンツの一部を「完全に捏造したテキスト」で回答した事例を公表している²⁾。

前兆はあった。2023年8月にNYTは自社のコンテンツに関する利用規約を改定し、AI開発のために記事・画像・動画などのあらゆるコンテンツをAIの学習データとして無断で利用することを禁止するとした。実は、この規約改定以前から、NYTはOpenAIを含む多くのAI開発企業と、自社コンテンツが学習データとして利用されていることに対して、その対価などについて協議を重ねてきたようだ。しかし合意に至らず、その結果が

今回の訴訟につながったとみられる。

NYTは訴状で、OpenAIとマイクロソフトに対して損害賠償と、自社コンテンツの学習データからの削除、AIモデルからの自社コンテンツの削除を要求している。

OpenAIは反論

一方で、OpenAIはNYTの訴訟に対して2024年1月8日、自社ブログに「NYTの訴訟には根拠がない」との反論記事をアップした³⁾。

OpenAIは、「NYTは完璧なコピーをChatGPTに回答させるよう誘導したり、NYTを騙るコンテンツを回答させたりするようなプロンプトを意図的に利用した可能性」を指摘し、加えて著作権で保護されているコンテンツであっても「フェアユース⁴⁾」の範囲内でのデータ活用は認められた権利であり、またAIの発展に不可欠なものだと主張している。

さらに、OpenAIのサム・アルトマンCEOは、「(AIの学習データに) 特定のデータソースが大きく意味を持つことはない。例えば高校レベルの生物学を理解するのに教科書を二千冊も読む必要はない」と述べ、AIの学習データにNYTといった特定のデータソースに依存する必要性はないと強調した⁵⁾。

訴訟の結末は不透明だが、AIの学習データに著作権で保護されているコンテンツを利用する際にどのように著作権を取り扱うべきなのかという点に関して、今回の裁判結果は非常に注目されている。

仮にNYTの主張が全面的に認められた場合、OpenAIは存続の危機にさらされるだろう。一方で、OpenAIの主張が認められるなら、コンテンツ産業のビ

NOTE

- 1) New York Times Sues OpenAI and Microsoft Over Use of Copyrighted Work - The New York Times <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>
- 2) The New York Times, MicrosoftとOpenAIをAIによる著作権侵害で提訴 - ITmedia NEWS <https://www.itmedia.co.jp/news/articles/2312/28/news100.html>
- 3) OpenAI and journalism <https://openai.com/blog/openai-and-journalism>
- 4) 「フェアユース」とは利用目的や引用形態などが一定の範囲にとどまっていれば著作権者の許諾を得ずとも著作物を利用できるという法理。日本では公正利用などと呼ばれる。アメリカなどでは著作権法で明文化されているが、日本では明文化された規定はない。
- 5) 米紙のデータ「学習に必要な」 オープンAIのトップ - 産経ニュース <https://www.sankei.com/article/20240119-EHRO3JZG75MP7OWBH44CCOFNLQ/>
- 6) 文化庁 文化審議会著作権分科会法制度小委員会 (第6回) より、資料1-1参照 (PDFファイル) https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_06/
- 7) 例えば日経225については以下を参照。日本経済新聞 ヘルプセンター https://www.nikkei.com/info/copyright_iphone.html
この中の「コンテンツの使用」には「日経コンテンツを許可なくデータマイニング、テキストマイニングおよびAI開発を目的としたディープラーニングなどの情報処理、情報解析のために、蓄積、複製、加工その他の利用を行うことはできません」との規定がある。

ジネスモデルは根本的な見直しが必要になる。

モデルの廃棄請求も可能になるとの見方を示している。

日本でもAIに関する著作権の議論が活発化

日本でもAIの学習データに関する著作権のあり方が議論されている。文化庁の文化審議会著作権分科会法制度小委員会は、「AIと著作権に関する考え方（素案）」

(以下、素案)を2024年1月15日に公開した⁶⁾。今回の素案は、「平成30著作権法改正」の際に定めたAIが著作物を利用するルールに関して、生成AIの登場に合わせて運用方針をアップデートしたものと見える。

今回提示された素案のポイントは、AIに関する著作権の論点について、「AIを学習・開発する段階」と「開発されたAIがコンテンツを生成・利用する段階」に分けて検討している点だ。現在の生成AIは、まず大量のデータを入力・分析してモデルを作成する(学習・開発段階)。そしてそこで構築されたAIモデルが、入力に応じた出力を行うようになっている(生成・利用段階)。

日本の著作権法では、学習・開発段階で利用する著作物の取り扱いは諸外国と比較して比較的緩やかな規定となっていると評価されてきた。しかし、今回の素案では、明らかに元の著作物を模倣する目的を持つ場合などでは学習段階でも著作権侵害を問える可能性を示した。

また、生成・利用段階についても、著作権侵害のおそれのあるAIモデルへの対抗措置に関してかなり踏み込んだ方針が示された。例えば、将来的にそのAIによる著作権侵害行為がかなりの蓋然性で予測されるときには、そのAIモデルから自社の学習データを削除するよう請求できる、いわゆる「予防措置」にも言及されている。さらに著作権侵害を高確率で起こすモデルに対してはそのモ

金融機関も他人事ではない

実は金融機関は多くの著作物を持つ著作権者だ。例えば日経225をはじめとする各種インデックスデータはその算出主体が一切の著作権と知的財産権を保持している⁷⁾。大量に発行しているレポートも著作物だ。これらの著作物に関する利用条件等は各社のガイドラインによって規定されており、そして多くの著作権保持者は無断でのAI学習への自社データの利用を禁じている。

ただ、冒頭で見たようにNYTも同様の規定を持っていたことを思い出すべきだろう。生成AIが無数の自社コンテンツを学習データとして利用することを個別にチェックできる可能性は低い。生成AIが引き起こしかねない事故として、ある金融機関が公表したアナリストレポートの要約をAIが生成した際に、当初のレポートと正反対の結論を「捏造」してしまうようなケースも考えられる。しかもそこに「出所」として当該金融機関名が記載されていた場合大きな問題になるだろう。

自社の著作物をきちんと保護し、必要であれば対価を求めること、また侵害された際の対抗措置を事前に講じておくことの必要性は高まっている。金融業界として早急に検討すべき課題であり、それぞれの金融機関も著作権法への感度を高めておくべきだ。

Writer's Profile



柏木 亮二 Ryoji Kashiwagi

金融ITインベーション事業本部
エキスパート研究員
専門はIT事業戦略分析
focus@nri.co.jp